Structure of LCP1 and alignments of domains

(729 a.a.) C (729 a.a.) C	cub LCCL DSD/FA58C	10 1 CGGTLTAS-S-G 26 CGHTVLGPES-G 9 CGRVIKDT-S-G 752 CEHKVTSTSG 484 CGGSFGGTQG 766 CGGFLFYA-SG 25 CGDTIKIL-SP-G 17 v£SAELLT-AhfG 520 CGGNMPGP-EG	#441 3 CGHTVLGP-ES-G "* "* "*
	FIG. 1A LCP1 (729 a.a	CUB d CUB d consens LCP1 1SPP_B gi 134! gi 1168 gi 5913 gi 5913 gi 2498	gi 70 1.0000 1.0000 1.0

10 20 30 40 50 60 70 80	QAVTCDTKALDLCKPVTDCNRVHCPAGCLLPKAKVFGTIVYASLSSICRAAVHAGVIDNTSGGAVDVVRVGGG1ITCLDTASNFLEPEFSKYCPAGCLLPKAEISGTIPHGYRDSSPLCMAGVHAGVVSNTLGGQISVVISKGGLITCLDTASNFLEPEFSKYCPAGCLLPKAEISGTIPHGYRDSSPLCMAGVHAGVVSNTLGGQISVVISKGGDLDCYTTVAQLCPFERPATHCPRIHCPAHCKDEPSywAPVFGTNIYADTSSICKTAVHAGVISNESGGDVDVMPVDKQIVSCEVRLRDQCKGTT-CNRYECPAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDND-GGWVDITRQGRQAVTCETTVEQLCPFRKPASHCPRVYCPRNCMQANPhYARVIGTRVYSDLSSICRAAIHYGIIDND-GGWVDITRQGRQAVRCDTKMKDRCKGST-CNRYQCPAGCLNHKAKIFGTLFYESSSSICRAAIHYGILDDK-GGLVDITRNGKNAITCFTRGLDIRKETEDVLCPANCPLWQFYVFGDGIYASLSSVCGAAVHRGVIGT-SGGPVRVYSLPGIPVTCFTRGLDIRKEKADVLCPAGCSLEEFSVFGNIVYASVSSICGAAVHRGVIGT-SGGPVRVYSLPGILFGHCIQFFWNLSSVSTAPAGCLLPFAEISGTIPHGYRDSSPLCMAGVVSNTLGGQISVVISKGRREVDCDSKAVDFLDDVGEPVRIHCPAGCSLTAATVVYGTDVYASYSSVCGAAVHRGVUDN-SGGAVHVVNNGPPQINCDVKAGKIIDPEFIVKCPAGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDN-SGGKILVRKVAG	90 * GSLANGIQSE 87 SSLANNVTSV 230 GGLANGIQTI 375 KSNRNGIQTI 375 ASFQNGIFSE 483 KSERHGVQSL 370 AVHANGIQSQ 108 SSUDANGIQSQ 114 SSLANNVTSV 208 GGSYSNGVQSL 124
10		90* * QPYYIGSLANGIQSE IPYYESSLANNVTSV KKTYVGSLRNGVQSE KHYFIKSNRNGIQTI RKTYIASFQNGIFSE VPFFVKSERHGVQSL QENYPAVHANGIQSQ RENYSSVDANGIQSQ IPYYESSLANNVTSV YSDFLGSDLNGIKSE QSGYKGSYSNGVQSL
FIG. 1C LCCL domain	consensus LCP1 gi 12053227 387 gi 12002311 291 gi 12002311 392 gi 12053227 286 gi 7387581 26 gi 7387581 26 gi 7387581 26 gi 7387581 26 gi 7387581 32	consensus LCP1 gi 12053227 gi 12002311 gi 12053227 gi 7387581 gi 7387581 gi 734441 gi 913964
	21	, 0

62 316 70 71 71 340 300 300 2093 295	133 133 134 144 145 145 146 136 136 136 136 136	
30 40 50 60 70 80 * * * * * * * * * * * * * * * *	160 LTWHG TQWQQ RTWNQ QSWVH QTWHL VAWHN VSWHN VSWHN VSWHN THYSI WSWYG	
70 WILGUT WILG	150 * IVARYIRILP IIARFIRVNP ILTRYLRIHP LLTRFIRIRP ILARYVRILP FMARYVRULP IIAQYIRLHP	
* NDONG NTDEYG NNNPKE NNNPKE SYGNDG SYGN	*. FDPPI FNPPI FETPI FEKPE FNPPI FEKPE FNPPI FEKPE	
60 XY Z XY Z XY Z XY Z XQ X XQ S XQ	140 .v. .vrind .vri	
* GAWSI CAWASI NAWQI NGWTI NAWVI INAWVI NAWWSI NAWWSI	*. DRNT) NTKG) NTKG) DSTT DATE) DNHS) DNNS DNNS	
50 ISQGGZ ISQGGR\ IRQGR\ ILQGR\ ILQGR\ IGDD IKSGS\ IYSGS\ IYSGS\ IYSGS\ IYSGS\ IXSGS\ IXSG\ IXSGS\ IXSG\ IXSG\ IXSGS\ IXSG\ IXXG\ IX	130 . FPGNSI FPGNKI FPGNKI FPGNNI	
* MARLIN KARLIN TRARLIN SYARLIN SYARLIN TIGRLI TIGRLI KKARLIN	SKPKII SKPKII SKVKVI SKNHKI SSSKI GSSKV RKDKV RKDKV	
30 40 50 6 * * * *YSANWTPENARLINSQGGAGAWSPK- GQENSWKPKKARLIKKPGPPWAAFKKSWWGdYWEPFRARLINAQGRVNAWQAKFTINMfatWSPSKARLHLQGRSNAWRQO FSDGrWTPQQSRLHGDDNGWTPNYKTW1llfsWNPSYARLDKQGRNFNAWVAGYKTW1lrafgWYPHLGRLDNQGKINAWTAQYGQWAPKLARLHYSGSINAWSTKGYGQWAPKLARLHYSGSINAWSTKGFRTLnmdmftwEPRKARLDKQGKVNAWTSGFRTLnmdmftwEPRKARLDKQGKVNAWTSG	120 120 140 15. *	
30 * -YSAN GQENS -FTNMfa -FTNMfa -FSDGr -YKTWGlhl: -YKTWGlhl: -YKTWGlhl: -YKTWGlhl: -YKTWGlhl:	···*· TWTTY EWKTYY EWKPY OWTLF OWMYY NWTEY NWTEY NWTEY CHESY	
20 30 DDSQITASSSYSAN ADPQITASSV1ewtdhtgQENS ENKQITASSYKKSWWG SDAQITASSYFTNMfa ANEQISASSYFTNMfa ANEQISASSYFSDGr- PDKQITASSSYKTWG1 RDFQITASSSYKTWG1 RDFQITASSSYKTWG1 RDFQITASSSYKTWG1 RDFQITASSSYKTWG1 RDFQITASSSYKTWG1 RDFQITASSYYKTWG1	110 SDDGR SDDGQ SEQGY SQDGH SQDGH STNGE STNGE SDDGV SDDGV	
ewtdh:	···*· YKVQY YRILY YTIHY YTIHY YTIHY YTIHY YKVAY YKVAY YKVAY YKVAY YKVAY	
20 PDSQITASSS ADPQITASSV1ew ENKQITASSF SDAQITASSF ANEQISASSF PDKQITASSY PDKQITASSY PDKQITASSY PDKQITASSY PDKQITASSS RDFQITASSS RDFQITASSS RDFQITASSY GSSNNSITVL	100 * QWVTSYKVQ YYVSAYRIL MYVKSYTIH MYVKEFLIS YYVKSYKLE QFVASYKLE LYVSQFIIM LYVSQFIIM LYVSQFIIM	
nain * * PLGMESG-RIPDSQITASSS ELGMESG-VIADPQITASSS PLGMESG-VIADPQITASSY- PLGMESG-RIBNQQITASSY- PLGMESG-RIANBQITASSY- PLGMESG-RIANBQISASST- PLGMESG-RIANBQISASS PLGMASG-RIANBSQITASSS PLGMASG-HIRDFQITASSS PLGMASG-HIRDFQITASSG PLGMASG-HIRDFQITASSY TLGMESGGDRGSSNNSITVL	*. GSS SEHN LITS LITS CHTQk-g GSV GHII GHY	
	90 . RRDFG STWVE CKSLS CKSLS VKSLL ALISRE ALISRE ARDFG	
main 10* PLGMESG-NI PLGMESG-VI PLGMESG-VI PLGMESG-VI PLGMESG-RI	90 CUITGEREDE GIITTGSTMV ALITTGCKSL GVTTQCKSL GVTTQCKSL GITTQGARNF GLITTQGARNF GLITTQGARDF GLITTQGARDF GLITTQGARDF GLITTGGARDF	 RI 139 KI 394 SI 148 QI 147 GI 418 RI 378 RI 454 RS 2171 RI 467 -L 362
C dom 1 P 249 t 5 P 280 P 233 P 311 P 322 P	63 63 63 63 63 63 63 63 63 63 63 63 63 6	138 H 3393 H 147 S 146 G 417 G 417 G 417 G 453 H 466 J 466 J
A58C sus 7128 5380 5493 2 8894	sus 7128 6380 058 5493 3894	5128 5380 5493 5493 5493 5493
FIG. 1D DSD/FA5 consensus LCP1 LCP1 LCP1 LCP2 M 1D7P M gi 254712 gi 250638 gi 264549 gi 264549 gi 264549 gi 264549 gi 264549 gi 264549	consensus LCP1 1CZT_A 1D7P_M gi 254712 gi 250638 gi 127058 gi 264549 gi 704441	consensus LCP1 1CZT_A 1D7P_M gi 254712 gi 250638 gi 127058 gi 264549
	£1/£	

Structure of the LCP gene (Chr. 3q12.1)

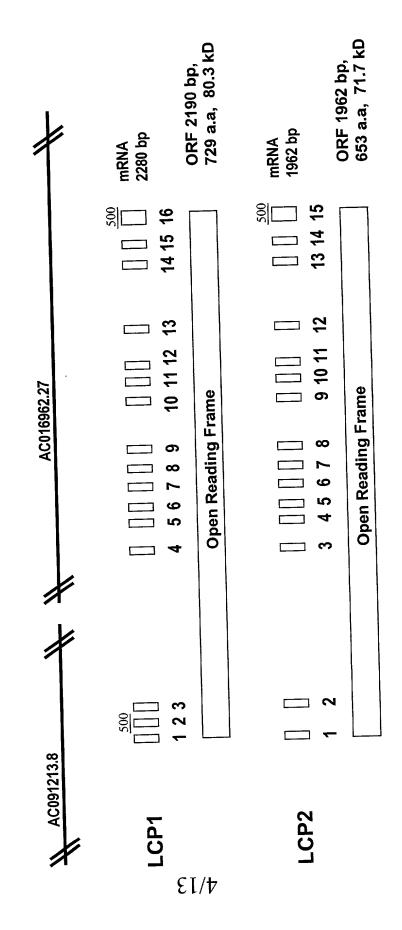


FIG. 2

LCP1

nt: SEQ ID NO: 1

aa: SEQ ID NO: 3

geogeogeoceogeotgggeogegetececetetecegetecetecet												52	
						М	P	L	F	L	L	L	7
gcto	caac	ctcct	ccto	ctto	ctcc	ATG			TTC	CTC	CTG	CTC	96
L	L	V	L	L	L	L	L	E	D	A	G	A	20
ATT	CTT	GTC	CTG	CTC	CTG	CTG	CTC	GAG	GAC	GCT	GGA	GCC	135
0	0	G	D	G	С	G	Н	т	v	L	G	P	33
			GAT								GGC	CCT	174
E	s	G	т	L	T	s	I	N	Y	P	Q	T	46
GAG	AGT	GGA	ACC	CTT	ACA	TCC	ATA	AAC	TAC	CCA	CAG	ACC	213
Y	P	N	s	т	V	С	E	W	E	I	R	V	59
TAT	CCC	AAC	AGC	ACT	GTT	TGT	GAA	TGG	GAG	ATC	CGT	GTA	252
K	M	G	E	R	V	R	I	K	F	G	D	F	72
AAG	ATG	GGA	GAG	AGA	GTT	CGC	ATC	AAA	TTT	GGT	GAC	TTT	291
D	I	E	D	ន	D	s	С	H	F	N	Y	L	85
GAC	ATT	GAA	GAT	TCT	GAT	TCT	TGT	CAC	TTT	AAT	TAC	TTG	330
R	r	Y	N	G	I	G	v	s	R	т	E	I	98
AGA	ATT	TAT	AAT	GGA	ATT	GGA	GTC	AGC	AGA	ACT	GAA	ATA	369
G	K	Y	C	G	L	G	L	Q	M	N	H	S	111
GGC	AAA	TAC	TGT	GGT	CTG	GGG	TTG	CAA	ATG	AAC	CAT	TCA	408
I	E	s	K	G	N	E	I	T	L	L	F	М	124
ATT	GAA	TCA	AAA	GGC	AAT	GAA	ATC	ACA	TTG	CTG	TTC	ATG	447
S	G	I	H	V	s	G	R	G	F	L	A	S	137
AGT	GGA	ATC	CAT	GTT	TCT	GGA	CGC	GGA	TTT	TTG	GCC	TCA	486
Y	s	V	I	D	K	Q	D	L	I	т	С	L	150
TAC	TCT	GTT	ATA	GAT	AAA	CAA	GAT	CTA	ATT	ACT	TGT	TTG	525
D	T	A	s	N	F	L	E	P	E	F	s	K	163
GAC	ACT	GCA	TCC	AAT	TTT	TTG	GAA	CCT	GAG	TTC	AGT	AAG	564

FIG. 3

Y С 176 ₽ Α G L L Ρ F Α E T TAC TGC CCA GCT GGT TGT CTG CTT CCT TTT GCT GAG ATA 603 Т I P Η G Y R D S S P 189 TCT GGA ACA ATT CCT CAT GGA TAT AGA GAT TCC TCG CCA 642 V Η 202 Α TTG TGC ATG GCT GGT GTG CAT GCA GGA GTA GTG TCA AAC 681 Т G G Q I S V V Ι S K G 215 ACG TTG GGC GGC CAA ATC AGT GTT GTA ATT AGT AAA GGT 720 Ţ Y Y E S S т 228 L Α Ν N ATT CCC TAT TAT GAA AGT TCT TTG GCT AAC AAC GTC ACA 759 V G S F T F Η L S T L 241 TCT GTG GTG GGA CAC TTA TCT ACA AGT CTT TTT ACA TTT 798 G C Y G T L Ε S 254 AAG ACA AGT GGA TGT TAT GGA ACA CTG GGG ATG GAG TCT 837 V 267 Т А D ₽ Q Ι \mathbf{T} Ά S S GGT GTG ATC GCG GAT CCT CAA ATA ACA GCA TCA TCT GTG 876 Η Т 280 G CTG GAG TGG ACT GAC CAC ACA GGG CAA GAG AAC AGT TGG 915 K K 293 Κ K А R L P AAA CCC AAA AAA GCC AGG CTG AAA AAA CCT GGA CCG CCT 954 A F Α \mathbf{T} D E Υ Q L Q 306 TGG GCT GCT TTT GCC ACT GAT GAA TAC CAG TGG TTA CAA 993 Т D L N K Ε K K Ι Τ G Ι Ι 319 ATA GAT TTG AAT AAG GAA AAG AAA ATA ACA GGC ATT ATA 1032 S G Т M V Ε Η N Y Υ V 332 ACC ACT GGA TCC ACC ATG GTG GAG CAC AAT TAC TAT GTG 1071 Υ R I L Y s D 345 TCT GCC TAC AGA ATC CTG TAC AGT GAT GAT GGG CAG AAA 1110 V γ R Ε Ρ G V Ε 0 D ĸ 358 TGG ACT GTG TAC AGA GAG CCT GGT GTG GAG CAA GAT AAG 1149 371 Ñ K D Y Η ATA TTT CAA GGA AAC AAA GAT TAT CAC CAG GAT GTG CGT 1188 P P Ι Ι Α R 384 AAT AAC TTT TTG CCA CCA ATT ATT GCA CGT TTT ATT AGA 1227

FIG. 3

V	N	P	T	Q	W	Q	Q	K	I	A	M ATG	K AAA	397 1266
												AAA	
M	E	L	L	G	C	Q	F	I	P	K	G	R	410 1305
ATG	GAG	CTG	CTC	GGA	TGT	CAG	.II.	A.I.I.	CCT	AAA	GGI	CGT	1303
P	Р	ĸ	L	T	Q	P	P	P	P	R	N	S	423
CCT	CCA	AAA	CTT	ACT	CAA	CCT	CCA	CCT	CCT	CGG	AAC	AGC	1344
Ŋ	D	T,	к	N	т	т	A	P	P	K	I	A	436
TAA	GAC	CTC	AAA	AAC	ACT	ACA	GCC	CCT	CCA	AAA	ATA	GCC	1383
7.7	C	מ	7\	D	ĸ	F	T	Ω	Þ	L	0	P	449
K AAA	GGT	CGT	GCC	CCA	AAA	TTT	ACG	CAA	CCA	CTA	CAA	CCT	1422
R	S ACT	S	N מידע מ	Ε	Triffic F	P	A GCA	Q CAG	ACA	GAA	CAA	T ACA	1461
T	A	S	P	D	I	R	N	T	T	V CTA	T act	P CCA	475 1500
N	V	Т	K	D	V	A	L	A	A	V	L	V	488
AAT	GTA	ACC	AAA	GAT	GTA	GCG	CTG	GCT	GCA	GTT	CTT	GTC	1539
P	V	L	v	М	V	L	\mathbf{T}	Т	L	I	L	I	501
CCT	GTG	CTG	GTC	ATG	GTC	CTC	ACT	ACT	CTC	ATT	CTC	ATA	1578
T,	V	С	A	W	н	W	R	N	R	K	K	K	514
TTA	GTG	TGT	GCT	TGG	CAC	TGG	AGA	AAC	AGA	AAG	AAA	AAA	1617
TT:		G	T	v	ח	т.	P	Y	W	D	R	A	527
ACT	GAA	. GGC	ACC	TAT	GAC	TTA	CCT	TAC	TGG	GAC	CGG	GCA	1656
													540
G GGT	W TGG	₩ TGG	Κ : ΑΔΑ	G GGA	M ATG	K AAG	Q CAG	r TTT	CTI	CCT	GCA	. AAA	1695
A	V	D . CAC	Н	E	E	T NCC	P CCA	V GTT	R ' CGC	צ דבי י	S AGC	S AGC	553 1734
GCA													
S	Ε	V	N	H	L	S	P	R	E	V	T	T	566 1773
AGC	GAA	A GTT	L'AA'	CAC	CTG	AGT	CCA	AGF	A GAA	ı Gic	. ACC	: ACA	1775
V	L	Q	Α	D	s	Α	E	Y	A	Q	P	L	579
GTG	CTC	G CAC	GCT	GAC	TCT	' GCA	GAG	; TA	r GC1	r CAG	; CCA	A CTG	1812
v	G	G	I	V	G	Т	L	Н	Q				592
GTA	GG	A GG2	rta <i>e</i>	GTI	GGT	ACA	CTT	CA	r car	A AGA	TC7	ACC	1851
F	K	P	F.	F.	G	K	Е	А	G	Y	A	D	605
TTI	' AA	A CC	A GAZ	A GAA	GGA	AAA	GAZ	A GCZ	A GG	C TAT	GC	A GAC	1890
												Н	
L CTA	D GA!	CC. ב	Y AT TA	и DAA C	S TCA	CCF	G GGO	G CA	G GA	A GTI	TAT	r CAT	

FIG. 3

A	Y	A	E	P	L	P	I	${f T}$	G	P	E	Y	631
GCC	TAT	GCT	GAA	CCA	CTC	CCA	ATT	ACG	GGG	CCT	GAG	TAT	1968
А	т	P	I	I	M	D	M	s	G	H	P	T	644
GCA	ACC	CCA	ATC	ATC	ATG	GAC	ATG	TCA	GGG	CAC	CCC	ACA	2007
т	S	V	G	0	P	s	${f T}$	s	Т	F	K	Α	657
ACT	TCA	GTT	GGT	CAG	CCC	TCC	ACA	TCC	ACT	TTC	AAG	GCT	2046
т	G	N	0	P	P	P	L	V	G	T	Y	N	670
ACG	GGG	AAC	CAA	CCT	CCC	CCA	CTA	GTG	GGA	ACT	TAC	AAT	2085
т	ь	L	s	R	Т	D	s	C	s	s	Α	Q	683
ACA	CTT	CTC	TCC	AGG	ACT	GAC	AGC	TGC	TCC	TCA	GCC	CAG	2124
А	0	Y	D	Т	P	K	Α	G	K	P	G	L	696
GCC	CAG	TAT	GAT	ACC	CCG	AAA	GCT	GGG	AAG	CCA	GGT	CTA	2163
P	Α	P	D	E	L	V	Y	Q	V	P	Q	S	709
CCT	GCC	CCA	GAC	GAA	TTG	GTG	TAC	CAG	GTG	CCA	CAG	AGC	2202
т	0	Е	V	s	G	А	G	R	D	G	E	C	722
ACA	CAA	GAA	GTA	TCA	GGA	GCA	GGA	AGG	GAT	GGG	GAA	TGT	2241
D	V	F	K	E	I	L	*						730
GAT	GTT	TTT	AAA	GAA	ATC	CTT	TGA	ag	atga	tgct	gctt	t	2280

LCP2

nt: SEQ ID NO: 1113 aa: SEQ ID NO: 1114

M	P	L	F	L	L	L	L	L	V	L	L	L	13
ATG	CCT	CTG	TTC	CTC	CTG	CTC	TTA	CTT	GTC	CTG	CTC	CTG	39
										K			26
CTG	CTC	GAG	GAC	GCT	GGA	GCC	CAG	CAA	GGC	AAA	TAC	TGT	78
										E			39
GGT	CTG	GGG	TTG	CAA	ATG	AAC	CAT	TCA	ATT	GAA	TCA	AAA	117
G		E								G			52
GGC	AAT	GAA	ATC	ACA	TTG	CTG	TTC	ATG	AGT	GGA	ATC	CAT	156
V	S	G	R	G	F	L	A	S	Y	S	V	I	65
GTT	TCT	GGA	CGC	GGA	TTT	TTG	GCC	TCA	TAC	TCT	GTT	ATA	195
D	K	Q	D	L	I	${f T}$	С	L	D	Т	A	S	78
GAT	AAA	CAA	GAT	CTA	ATT	ACT	TGT	TTG	GAC	ACT	GCA	TCC	234
												A	
AAT	TTT	TTG	GAA	CCT	GAG	TTC	AGT	AAG	TAC	TGC	CCA	GCT	273
												I	
GGT	TGT	CTG	CTT	CCT	TTT	GCT	GAG	ATA	TCT	GGA	ACA	ATT	312
P												Α	
CCT	CAT	GGA	TAT	AGA	GAT	TCC	TCG	CCA	TTG	TGC	ATG	GCT	351
												G	130
GGT	GTG	CAT	GCA	GGA	GTA	GTG	TCA	AAC	ACG	TTG	GGC	GGC	390
												Y	
CAA	ATC	AGT	GTT	GTA	ATT	AGT	AAA	GGT	ATT	· CCC	TAT	TAT	429
E												G	
GAA	AGT	TCT	TTG	GCT	AAC	AAC	GTC	ACA	TCT	GTG	GTG	GGA	468

FIG. 4

H	L	S	${f T}$	S	L	F	\mathbf{T}	F	K	T	S	G	169
CAC	TTA	TCT	ACA	AGT	CTT	TTT	ACA	TTT	AAG	ACA	AGT	GGA	507
~	37	a		7	a	3.6	_	0	~	7.7	-	~	100
			T										182
161	IAI	GGA	ACA	CIG	טטט	AIG	GAG	ICI	GGT	GIG	ATC	UUU	546
D	P	0	I	\mathbf{T}	Α	S	S	V	L	E	W	Т	195
			ATA										585
D	H	Т	G	Q	E	N	S	W	K	P	K	K	208
GAC	CAC	ACA	GGG	CAA	GAG	AAC	AGT	TGG	AAA	CCC	AAA	AAA	624
7	_	-			_	~	_	_		_	_	_	
			K										221
GCC	AGG	CIG	AAA	AAA	CCT	GGA	CCG	CCI	166	GCT	GCT	1.1.1	663
Α	Т	D	E	Y	0	W	L	0	Т	D	T,	N	234
												AAT	
K	E	K	K	I	T	G	I	I	${f T}$	${f T}$	G	S	247
AAG	GAA	AAG	AAA	ATA	ACA	GGC	ATT	ATA	ACC	ACT	GGA	TCC	741
т.	ъr	7.7	-), T	37	37	7.7	_	70	37	Б	260
			E										260
			E GAG										260 780
ACC	ATG	GTG	GAG	CAC	AAT	TAC	TAT	GTG	TCT	GCC	TAC	AGA	780
ACC I	ATG L	GTG Y	GAG	CAC D	AAT D	TAC G	TAT	GTG K	TCT W	GCC T	TAC V	AGA Y	780
ACC I	ATG L	GTG Y	GAG S	CAC D	AAT D	TAC G	TAT	GTG K	TCT W	GCC T	TAC V	AGA Y	780 273
ACC I ATC R	ATG L CTG E	GTG Y TAC P	GAG S AGT G	CAC D GAT V	AAT D GAT	TAC G GGG	TAT Q CAG D	GTG K AAA K	TCT W TGG	GCC T ACT F	TAC V GTG	AGA Y TAC G	780 273 819 286
ACC I ATC R	ATG L CTG E	GTG Y TAC P	GAG S AGT G	CAC D GAT V	AAT D GAT	TAC G GGG	TAT Q CAG D	GTG K AAA K	TCT W TGG	GCC T ACT F	TAC V GTG	AGA Y TAC	780 273 819 286
ACC I ATC R AGA	L CTG E GAG	GTG Y TAC P CCT	GAG S AGT G GGT	CAC D GAT V GTG	D GAT E GAG	G GGG Q CAA	Q CAG D GAT	GTG K AAA K AAG	TCT W TGG I ATA	GCC T ACT F TTT	V GTG Q CAA	Y TAC G GGA	780 273 819 286 858
ACC I ATC R AGA	ATG L CTG E GAG	GTG Y TAC P CCT	GAG S AGT G GGT	CAC DGAT VGTG	AAT DGAT EGAG	TAC GGGG QCAA D	TAT Q CAG D GAT V	GTG K AAA K AAG	TCT W TGG I ATA	GCC TACT FTTT	TAC V GTG Q CAA F	AGA Y TAC G GGA	780 273 819 286 858
ACC I ATC R AGA	ATG L CTG E GAG	GTG Y TAC P CCT	GAG S AGT G GGT	CAC DGAT VGTG	AAT DGAT EGAG	TAC GGGG QCAA D	TAT Q CAG D GAT V	GTG K AAA K AAG	TCT W TGG I ATA	GCC TACT FTTT	TAC V GTG Q CAA F	Y TAC G GGA	780 273 819 286 858
ACC I ATC R AGA N AAC	L CTG E GAG K AAA	YTAC PCCT DGAT	GAG S AGT G GGT Y TAT	CAC D GAT V GTG H CAC	D GAT E GAG Q CAG	G GGG Q CAA D GAT	Q CAG D GAT V GTG	GTG K AAA K AAG R CGT	TCT W TGG I ATA N AAT	GCC T ACT F TTT N AAC	V GTG Q CAA F TTT	AGA Y TAC G GGA	780 273 819 286 858 299 897
ACC I ATC R AGA N AAC	L CTG E GAG K AAA	GTG Y TAC P CCT D GAT	GAG S AGT G GGT Y TAT	CAC DGAT VGTG HCAC	D GAT E GAG CAG	GGGG Q CAA D GAT	Q CAG D GAT V GTG	GTG K AAA K AAG R CGT R	TCT W TGG I ATA N AAT	GCC T ACT F TTT N AAC	V GTG Q CAA F TTT	AGA Y TAC G GGA L TTG	780 273 819 286 858 299 897
ACC I ATC R AGA N AAC	L CTG E GAG K AAA	GTG Y TAC P CCT D GAT	GAG S AGT G GGT Y TAT	CAC DGAT VGTG HCAC	D GAT E GAG CAG	GGGG Q CAA D GAT	Q CAG D GAT V GTG	GTG K AAA K AAG R CGT R	TCT W TGG I ATA N AAT	GCC T ACT F TTT N AAC	V GTG Q CAA F TTT	AGA Y TAC G GGA L TTG	780 273 819 286 858 299 897
ACC I ATC R AGA N AAC P CCA	L CTG E GAG K AAA P CCA	YTAC PCCT DGAT IATT	GAG S AGT G GGT Y TAT I ATT	D GAT V GTG H CAC	D GAT E GAG CAG CAG	GGGG Q CAA D GAT F TTT A	Q CAG D GAT V GTG I ATT	KAAA KAAG R CGT R AGA	W TGG I ATA N AAT V GTG M	GCC T ACT F TTT N AAC N AAT	V GTG Q CAA F TTT P CCT	AGA Y TAC G GGA L TTG T ACC	780 273 819 286 858 299 897 312 936
ACC I ATC R AGA N AAC P CCA	L CTG E GAG K AAA P CCA	YTAC PCCT DGAT IATT	GAG S AGT G GGT Y TAT I ATT	D GAT V GTG H CAC	D GAT E GAG CAG CAG	GGGG Q CAA D GAT F TTT A	Q CAG D GAT V GTG I ATT	KAAA KAAG R CGT R AGA	W TGG I ATA N AAT V GTG M	GCC T ACT F TTT N AAC N AAT	V GTG Q CAA F TTT P CCT	AGA Y TAC G GGA L TTG T ACC	780 273 819 286 858 299 897 312 936
ACC I ATC R AGA N AAC P CCA Q CAA	L CTG E GAG K AAA P CCA W	YTAC PCCT DGAT IATT QCAG	GAG S AGT G GGT Y TAT I ATT Q CAG	D GAT V GTG H CAC A GCA	D GAT E GAG CAG CAG I ATT	GGGG Q CAA D GAT F TTT A GCC	Q CAG D GAT V GTG I ATT	KAAA KAAG RCGT RAGA KAAA	W TGG I ATA N AAT V GTG M ATG	GCC T ACT F TTT N AAC N AAT E GAG	V GTG Q CAA F TTT P CCT	AGA Y TAC G GGA L TTG T ACC L CTC	780 273 819 286 858 299 897 312 936 325 975
ACC I ATC R AGA N AAC P CCA Q CAA	L CTG E GAG K AAA P CCA W TGG	YTAC PCCT DGAT IATT QCAG	GAG S AGT G GGT Y TAT I ATT Q CAG	CAC D GAT V GTG H CAC A GCA K AAA	D GAT E GAG CAG CAG ATT	GGGG Q CAA D GAT F TTT A GCC K	Q CAG D GAT V GTG I ATT M ATG	KAAA KAAG RCGT RAGA KAAA	W TGG I ATA N AAT V GTG M ATG	GCC T ACT F TTT N AAC N AAT E GAG	V GTG Q CAA F TTT P CCT L CTG	AGA Y TAC G GGA L TTG T ACC L CTC	780 273 819 286 858 299 897 312 936

FIG. 4

T	Q	P	P	Р	Р	R	N	S	N	D	L CTC	K aaa	351 1053
ACI	CAA	CCI	CCA	CCI	CCI	CGG	AAC	AGC	AAI	Oric	CIC	1001	1033
N	T	${f T}$	A	P	P	K	I	A	K	G	R	A	
AAC	ACT	ACA	GCC	CCT	CCA	AAA	ATA	GCC	AAA	GGT	CGT	GCC	1092
P	K	F	Т	Q	P	L	Q	P	R	S	S	N	377
CCA	AAA	TTT	ACG	CAA	CCA	CTA	CAA	CCT	CGC	AGT	AGC	AAT	1131
E	F	P	A	Q	Т	E	Q	${f T}$	Т	A	s	P	390
GAA	TTT	CCT	GCA	CAG	ACA	GAA	CAA	ACA	ACT	GCC	AGT	CCT	1170
D													403
GAT	ATC	AGA	AAT	ACT	ACC	GTA	ACT	CCA	AAT	GTA	ACC	AAA	1209
D	V	Α	L	A	A	V	L	V	P	V	L	V	416
GAT	GTA	GCG	CTG	GCT	GCA	GTT	CTT	GTC	CCT	GTG	CTG	GTC	1248
М	V	L	Т	Т	L	I	L	I	L	V	С	A	429
ATG	GTC	CTC	ACT	ACT	CTC	ATT	CTC	ATA	TTA	GTG	TGT	GCT	1287
												T	
TGG	CAC	TGG	AGA	AAC	AGA	AAG	AAA	AAA	ACT	GAA	GGC	ACC	1326
													455
TAT	GAC	TTA	CCT	TAC	TGG	GAC	CGG	GCA	GGT	TGG	TGG	AAA	1365
G	М	K	0	F	L	P	A	K	A	V	D	H	468
GGA	ATG	AAG	CAG	TTT	CTT	CCT	GCA	AAA	GCA	GTG	GAC	CAT	1404
E	E	${f T}$	P	V	R	Y	S	S	S	E	V	N	481
GAG	GAA	ACC	CCA	GTT	CGC	TAT	AGC	AGC	AGC	GAA	GTT	AAT	1443
Н	L	S	P	R	E	V	Т	Т	V	L	Q	A	494
CAC	CTG	AGT	CCA	AGA	GAA	GTC	ACC	ACA	GTG	CTG	CAG	GCT	1482
D	S	Α	E	Y	A	Q	P	L	V	G	G	I	507
GAC	TCT	GCA	GAG	TAT	GCT	CAG	CCA	CTG	GTA	GGA	GGA	ATT	1521
V	G	Т	L	Н	Q	R	S	${f T}$	F	K	P	E	520
GTT	GGT	ACA	CTT	CAT	CAA	AGA	TCT	ACC	TTT	AAA	CCA	GAA	1560
E	G	K	E	A	G	Y	A	D	L	D	P	Y	533

FIG. 4

GAA	GGA	AAA	GAA	GCA	GGC	TAT	GCA	GAC	CTA	GAT	CCT	TAC	1599
N	S	P	G	Q	E	V	Y	Н	A	Y	А	E	546
AAC	TCA	CCA	GGG	CAG	GAA	GTT	TAT	CAT	GCC	TAT	GCT	GAA	1638
													559
CCA	CTC	CCA	ATT	ACG	GGG	CCT	GAG	TAT	GCA	ACC	CCA	ATC	1677
													572
ATC	ATG	GAC	ATG	TCA	GGG	CAC	CCC	ACA	ACT	TCA	GTT	GGT	1716
Q	P	S	Т	S	Т	F	K	Α	Т	G	N	Q	585
CAG	CCC	TCC	ACA	TCC	ACT	TTC	AAG	GCT	ACG	GGG	AAC	CAA	1755
P	P	P	L	V	G	${f T}$	Y	N	Т	L	L	S	598
CCT	CCC	CCA	CTA	GTG	GGA	ACT	TAC	AAT	ACA	CTT	CTC	TCC	1794
R	\mathbf{T}	D	S	С	S	S	A	Q	A	Q	Y	D	611
AGG	ACT	GAC	AGC	TGC	TCC	TCA	GCC	CAG	GCC	CAG	TAT	GAT	1833
													624
ACC	CCG	AAA	GCT	GGG	AAG	CCA	GGT	CTA	CCT	GCC	CCA	GAC	1872
													637
GAA	TTG	GTG	TAC	CAG	GTG	CCA	CAG	AGC	ACA	CAA	GAA	GTA	1911
													650
TCA	GGA	GCA	GGA	AGG	GAT	GGG	GAA	TGT	GAT	GTT	TTT	AAA	1950
		L											654
GAA	ATC	CTT	TGA										1962

200 bp